Finite Element Solution of Optimal Control Problems Arising in Semiconductor Modeling

Pavel Bochev and Denis Ridzal

Computational Mathematics and Algorithms Department,
Sandia National Laboratories[1], Albuquerque, NM 87185-1320, USA,
{pbboche,dridzal}@sandia.gov

Abstract. Optimal design, parameter estimation, and inverse problems arising in the modeling of semiconductor devices lead to optimization problems constrained by systems of PDEs. We study the impact of different state equation discretizations on optimization problems whose objective functionals involve flux terms. Galerkin methods, in which the flux is a derived quantity, are compared with mixed Galerkin discretizations where the flux is approximated directly. Our results show that the latter approach leads to more robust and accurate solutions of the optimization problem, especially for highly heterogeneous materials with large jumps in material properties.

1 Introduction

Common objectives in the modeling of semiconductor devices are, e.g., to control the current flow over a contact by changing the so called doping profile of the device (optimal design problem), or to characterize an unknown doping profile based on measurements of the current flow (inverse problem).

In either case, the resulting PDE constrained optimization problems call for objective functionals that involve flux terms in their definition. Depending on whether the state equation is discretized by a Galerkin or a mixed Galerkin method, flux terms can have fundamentally different representations. For instance, the Galerkin method approximates the scalar concentration variables by finite element subspaces of $H^1(\Omega)$, and the flux is a derived quantity, while in the mixed method the flux is approximated directly by subspaces of $H(\mathrm{div}, \Omega)$.

While numerical solution of optimization problems arising in semiconductor device modeling has been previously addressed in the literature [1, 2], there are virtually no studies on how discretization choices impact the accuracy and the robustness of the numerical approximation. To a degree, our work is motivated by earlier studies [3, 4] of optimization problems governed by advection-dominated PDEs, which showed that stabilization of the state equations and stabilization of the optimality system yield different solutions of the optimization problem.

[1] Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94-AL85000.

However, in contrast to these papers, our main focus is on how different discrete formulations of the state equation may affect the optimization problem.

Our report is organized as follows. The model optimization problem is described in Section 2. Section 3 states the Galerkin and mixed Galerkin discretizations of the optimization problem (1). Numerical results contrasting these methods, and a discussion of the results, are presented in Section 4.

2 Model Optimization Problem

A common objective in the design of semiconductor devices is to match the current J measured at a portion Γ_o of the Dirichlet boundary (see Fig. 1) to a prescribed value $\widehat{J}$, while allowing for "small" (controlled) deviations of the doping profile u from a reference doping profile $\widehat{u}$. For a complete formulation of such optimization problems, constrained by the drift-diffusion semiconductor equations, we refer to [1, 2, 5].

The primary goal of this work is to study how different discretizations of the state equations impact the solution of the optimization problem. This question can be investigated on a much simpler model, and so, we restrict our attention to the following linear–quadratic elliptic optimization[1] problem,

$$\text{minimize} \quad \frac{1}{2}\|\nabla y \cdot \nu - \nabla \widehat{y} \cdot \nu\|^2_{-1/2,\Gamma_o} + \frac{\alpha}{2}\|u - \widehat{u}\|^2_{0,\Omega} \tag{1a}$$

subject to

$$-\nabla \cdot (k(x)\nabla y(x)) = f(x) + u(x) \qquad \text{in } \Omega \tag{1b}$$

$$y(x) = y_D(x) \qquad \text{on } \Gamma_D \tag{1c}$$

$$(k(x)\nabla y(x)) \cdot \nu = g(x) \qquad \text{on } \Gamma_N, \tag{1d}$$

where $\alpha, k(x) > 0$, $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$ is a bounded domain, and Γ_D, Γ_N are the Dirichlet and Neumann parts of $\partial\Omega$. We assume that $\Gamma_D \neq \emptyset$ and $\Gamma_N = \partial\Omega/\Gamma_D$. Model (1) follows from the full problem considered in [5] by assuming that the electron and hole densities are given functions.

3 Discretization of the Optimization Problem

We consider two discretizations of (1) that differ by their choice of finite element methods for the state equation. In each case we discuss computation of the flux terms in the objective functional, necessary to complete the discretization of the optimization problem.

For more details regarding the solution of the discrete optimization problem, or the existence and uniqueness of optimal solutions of (1) we refer to [1, 2, 5].

[1] As it is customary in the optimal control context, we refer to the variables $y(x)$ in (1) as the *state variables*, whereas the doping profile $u(x)$ will play the role of the *control variables*. Equation (1b) is known as the *state equation*. Additionally, we will often abbreviate $y(x)$, $u(x)$, etc., by y, u, etc.

3.1 Galerkin Discretization

We define the state and control spaces $Y = \left\{ y \in H^1(\Omega) \,:\, y = y_D \text{ on } \Gamma_D \right\}$, $U = L^2(\Omega)$, and the space of test functions $V_0 = \left\{ v \in H^1(\Omega) \,:\, v = 0 \text{ on } \Gamma_D \right\}$. The weak form of (1) is to find $y \in Y, u \in U$, which solve the problem

$$\text{minimize} \quad \frac{1}{2}\|\nabla y \cdot \nu - \nabla \widehat{y} \cdot \nu\|^2_{-1/2,\Gamma_o} + \frac{\alpha}{2}\|u - \widehat{u}\|^2_{0,\Omega} \tag{2a}$$

subject to

$$a(y,v) + b(u,v) = (f,v) + \langle g, v \rangle_{\Gamma_N}, \quad \forall v \in V_0 \tag{2b}$$

where $\langle \cdot, \cdot \rangle_\star$ denotes the duality pairing between $H^{-1/2}(\star)$ and $H^{1/2}(\star)$, and

$$a(y,v) = \int_\Omega k \nabla y \cdot \nabla v \, dx, \qquad b(u,v) = -\int_\Omega uv \, dx, \qquad (f,v) = \int_\Omega fv \, dx \,.$$

The finite element discretization of the state equation is obtained in the usual manner by restricting (2b) to finite element subspaces $Y_h \subset Y$, $V_{0,h} \subset V_0$ and $U_h \subset U$ of the state, test, and control spaces, respectively.

When using a Galerkin method for the state equation, discretization of the objective functional (2a) requires additional attention, because the flux $\nabla y \cdot \ \nu$ appearing in (2a) is not approximated directly by the method. A standard approach to discretizing the flux term $\|\nabla y \cdot \nu - \nabla \widehat{y} \cdot \nu\|^2_{-1/2,\Gamma_o}$ would be to restrict it to the finite element space Y_h for the states, and then use a weighted L^2 norm to approximate the norm on $H^{-1/2}(\Gamma_o)$:

$$\|\nabla y \cdot \nu - \nabla \widehat{y} \cdot \nu\|^2_{-1/2,\Gamma_o} \approx h \|\nabla y_h \cdot \nu - \nabla \widehat{y}_h \cdot \nu\|^2_{0,\Gamma_o} \,. \tag{3}$$

While this discretization of the flux term is consistent, in the sense that every instance of the state y in the optimization problem is approximated by the same finite element basis, it may not be the best possible choice for this term. This is certainly true if the state equation is solved separately and ∇y_h is used to approximate the flux.

It is well-known that a more accurate flux approximation can be obtained by postprocessing the finite element solution, instead of simply taking its derivative. One such technique is the variational flux approximation (VFA) [6–9]. It is based on the Green's formula and has been applied to optimization problems by Berggren et. al. in [10]. In the VFA approach, the standard flux $\nabla y_h \cdot \nu$ is replaced by a more accurate, C^0 approximation λ_h, obtained by solving the equation

$$\int_{\Gamma_0} \lambda_h v_h \, dl = k^{-1} \left(a(y_h, v_h) + b(u_h, v_h) - (f, v_h) - \int_{\Gamma \setminus \Gamma_o} k \nabla y_h \cdot \nu v_h \, dl \right) . \tag{4}$$

Using VFA we approximate the flux term as follows:

$$\|\nabla y \cdot \nu - \nabla \widehat{y} \cdot \nu\|^2_{-1/2,\Gamma_o} \approx h \|\lambda_h - \widehat{\lambda}_h\|^2_{0,\Gamma_o} \,, \tag{5}$$

For implementation details of VFA we refer to the above papers.

Remark 1. The use of VFA in our optimization problem differs substantially from its use as a postprocessing technique and in the optimization problem of [10]. When (4) is used to postprocess a given finite element solution (y_h, u_h), the right hand side in (4) involves only known quantities. In [10], VFA is used in an already defined optimality system to improve the accuracy of the solution. In contrast, in our case, the VFA approximation *changes* the optimization problem, because the discretization of $\|\nabla y \cdot \nu - \nabla \widehat{y} \cdot \nu\|^2_{-1/2,\Gamma_o}$ by (5) makes this term a function of *both* the unknown state y_h and the unknown control u_h.

3.2 Mixed Galerkin Discretization

A mixed Galerkin method for the state equation is defined by using its equivalent first-order system form

$$\begin{cases} \nabla \cdot p + u = -f \text{ in } \Omega \\ k^{-1}\, p - \nabla y \;\; = 0 \;\; \text{in } \Omega \end{cases} \quad \text{and} \quad \begin{aligned} y &= y_D \text{ on } \Gamma_D \\ (k\nabla y)\cdot \nu \;\; &= g \;\; \text{on } \Gamma_N. \end{aligned} \tag{6}$$

For the variational formulation of the optimization problem, we introduce[2] the state spaces $Y = L^2(\Omega)$ and $P = H_{g,N}(\mathrm{div}, \Omega)$, the control space $U = L^2(\Omega)$, and the trial spaces $V = L^2(\Omega)$ and $Q_0 = H_{0,N}(\mathrm{div}, \Omega)$. To simplify the notation we write $\widehat{p} = k\, \nabla \widehat{y}$. The weak form of (1), using the mixed Galerkin discretization of (6), is to find $y \in Y, p \in P, u \in U$, which solve the problem

$$\text{minimize} \quad \frac{1}{2}\|k^{-2}(p\cdot\nu - \widehat{p}\cdot\nu)\|^2_{-1/2,\Gamma_o} + \frac{\alpha}{2}\|u - \widehat{u}\|^2_{0,\Omega}, \tag{7a}$$

subject to

$$\begin{aligned} a(p,q) + b(q,y) &= \langle y_d, q\cdot\nu\rangle_{\Gamma_D} \qquad \forall q \in Q_0 \\ b(p,v) + c(u,v) &= -\,(f,v) \qquad\qquad \forall v \in V, \end{aligned} \tag{7b}$$

where $(\cdot,\cdot)$, and $\langle\cdot,\cdot\rangle_\star$ were defined in Sec. 3.1, and

$$a(p,q) = \int_\Omega k^{-1}\, p\cdot q\; dx, \qquad b(q,y) = \int_\Omega (\nabla\cdot q) y\; dx, \qquad c(u,v) = \int_\Omega uv\; dx.$$

The mixed finite element discretization of the state equation follows by restricting (7b) to finite element subspaces $Y_h \subset Y$, $P_h \subset P$, $U_h \subset U$, $V_h \subset V$, and $Q_{0,h} \subset Q_0$, for the states, controls, and the respective test functions in (7b). We recall that the pairs (Y_h, P_h) and $(Q_{0,h}, V_h)$ are subject to an inf-sup stability condition [11].

In contrast to the Galerkin approach in Sec. 3.1, in the mixed method the flux is approximated directly by p_h. As a result, the flux term in the objective functional can be discretized as follows:

$$\|k^{-2}(p\cdot\nu - \widehat{p}\cdot\nu)\|^2_{-1/2,\Gamma_o} \approx h\|k^{-2}(p_h\cdot\nu - \widehat{p}_h\cdot\nu\|^2_{0,\Gamma_o}\,. \tag{8}$$

[2] We recall that $H(\mathrm{div},\Omega) = \left\{q \in \left[L^2(\Omega)\right]^2 : \nabla\cdot q \in L^2(\Omega)\right\}$; $H_{0,N}(\mathrm{div},\Omega)$ is the subspace of all fields in $H(\mathrm{div},\Omega)$ whose normal component vanishes on Γ_N, and $H_{g,N}(\mathrm{div},\Omega)$ are the fields in $H(\mathrm{div},\Omega)$ whose normal component on Γ_N equals g.

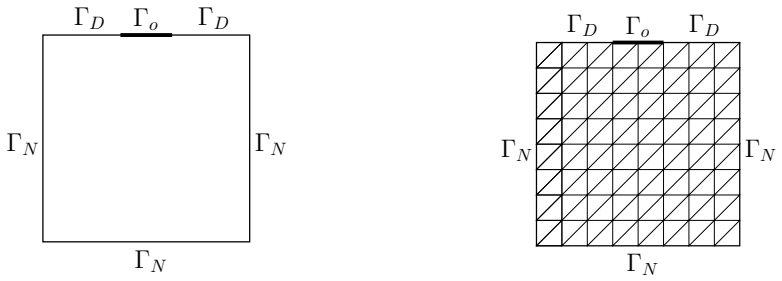

Fig. 1. Computational domain (left) and its partition into finite elements (right plot).

	Example 1			Example 2		
	GM	Mixed	GM-VFA	GM	Mixed	GM-VFA
$\mathcal{J}_F$	1.99e-06	1.88e-06	1.95e-04	6.42e-09	2.51e-09	2.70e-09
$\mathcal{J}_u$	1.10e-08	1.10e-08	3.81e-07	1.12e-03	1.08e-03	1.11e-03
$\mathcal{J}$	2.00e-06	1.89e-06	1.95e-04	1.12e-03	1.08e-03	1.11e-03
	Example 3			Example 4		
	GM	Mixed	GM-VFA	GM	Mixed	GM-VFA
$\mathcal{J}_F$	6.07e-05	8.10e-10	2.83e-07	1.06e+01	2.56e-07	6.29e-07
$\mathcal{J}_u$	7.17e-05	4.62e-05	4.50e-03	3.63e+00	4.57e-03	7.19e-03
$\mathcal{J}$	1.32e-04	4.62e-05	4.50e-03	1.42e+01	4.57e-03	7.19e-03

Table 1. $\mathcal{J}_F$, $\mathcal{J}_u$ and $\mathcal{J}$ denote the values of the flux term, the control term and their sum (the total value of the objective functional).

4 Numerical Results

The computational domain $\Omega = [-1, 1]^2$, its finite element partition, and the boundary part Γ_o are shown in Fig. 1. All numerical results were obtained on a 32×32 mesh with 2048 triangles, 1082 vertices, and 3136 edges. For the Galerkin method in Sec. 3.1, we use standard C^0, piecewise linear approximation spaces. The mixed Galerkin method for the state equation was implemented using the lowest order Raviart-Thomas element for p_h and piecewise constant finite elements for y_h.

In the examples below we compare finite element solutions of the optimization problem using the standard Galerkin method, the Galerkin method with VFA, and the mixed method. The most distinguishing characteristic of the examples used in our study are the differences in the corresponding diffusivity (i.e. permittivity, in the case of semiconductors) profiles $k(x)$. For all examples we use $f(x) = 0$, $y_D = 0$, $\widehat{u} = 1$, and $\alpha = 6.25 \cdot 10^{-4}$. The Neumann data g is set to 0 on the left and right sides of Ω and to $-k(x)$ on the bottom.

Example 1. The desired flux is $\nabla \widehat{y} \cdot \nu = 1$ and $k(x) = 10^2$ in Ω.

Example 2. The desired flux is $\nabla\widehat{y}\cdot\nu = 1$ and $k(x) = 10^{-2}$ in Ω.
Example 3. The desired flux is $\nabla\widehat{y}\cdot\nu = 1$ and

$$k(x) = \begin{cases} 10 & \text{in } [-1,-0.25]\times[-1,1] \\ 10^{-2} & \text{in } [-0.25,0.25]\times[-1,1] \\ 10 & \text{in } [0.25,1]\times[-1,1], \end{cases}$$

Example 4. The desired flux is $\nabla\widehat{y}\cdot\nu = 100$ and $k(x)$ is as in Ex. 3.

Objective functional values for the four examples are summarized in Table 1. The data for Ex. 1-2 shows that for constant $k(x)$ all three discretizations perform at a comparable level. Nevertheless, the mixed Galerkin method does consistently outperform the other two discretizations, albeit by a small margin. We also observe that the VFA approach gives better results (i.e., closer to the mixed Galerkin results) for $k(x) \ll 1$, while the standard flux approximation does better for $k(x) \gg 1$.

The data for Ex. 3-4 shows a completely different situation, as far as the standard flux approximation is concerned. For these two examples the mixed Galerkin method clearly outperforms the standard Galerkin discretization of the optimization problem, especially when the desired flux value is large, as in Ex. 4. Using the VFA approach, the Galerkin discretization fares better, how-

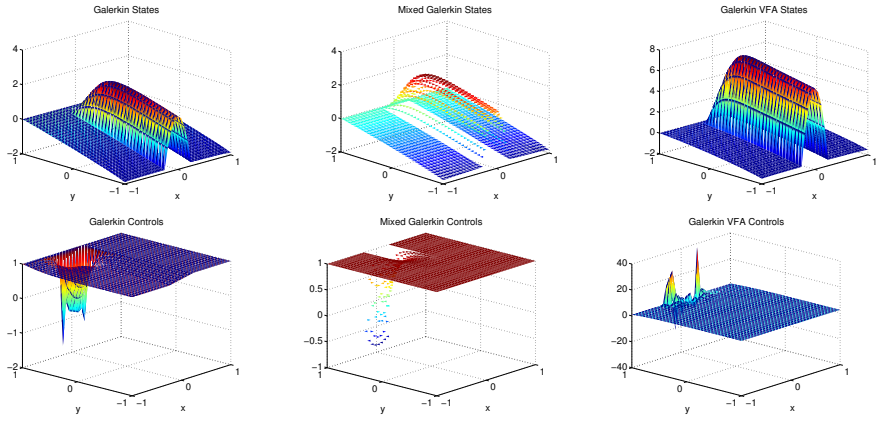

Fig. 2. Galerkin, mixed Galerkin, and Galerkin VFA optimal states (top row) and optimal controls (bottom row) for Ex. 3.

ever, the objective functional values remain less accurate than those computed with the mixed method. These observations are also confirmed by the plots in Fig. 2-3. We see that for Example 4 the states and controls computed by the standard Galerkin method are grossly inaccurate. We also note that among all three methods the controls computed by the mixed method exhibit the most

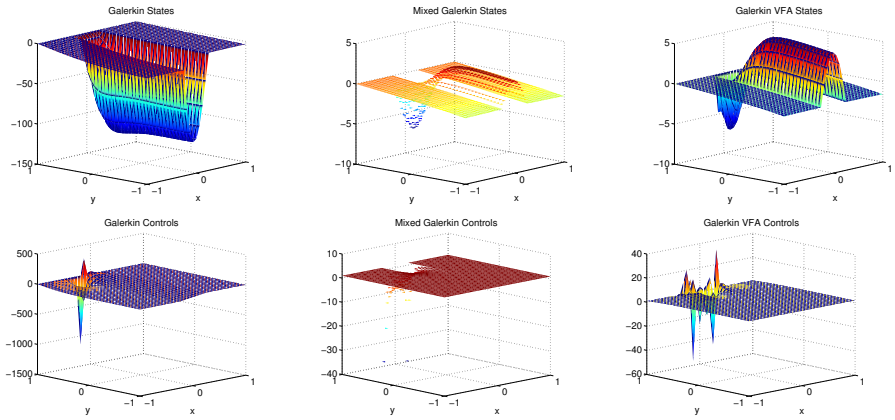

Fig. 3. Galerkin, mixed Galerkin, and Galerkin VFA optimal states (top row) and optimal controls (bottom row) for Ex. 4.

robust behavior. An interesting feature of the controls for the VFA approach is their oscillatory nature. This could be a problem in some specific applications where the controls have to be implemented in real materials.

Based on the numerical data, we can conclude that for problems with heterogeneous material properties the mixed Galerkin method offers the most robust performance and the most accurate results. The worst performer is the standard Galerkin method, which may yield state and control approximations that are many orders of magnitude less accurate than those computed by the mixed method. Thus, we cannot recommend the standard Galerkin discretization as a reliable approach to solving optimization problems whose objective functionals involve flux terms. Instead, for such problems, one should use the mixed Galerkin discretization whenever possible. If, for whatever reason, the use of the mixed method is not feasible, then the Galerkin discretization of the state equations should be combined with the VFA approach in order to improve robustness and accuracy.

Acknowledgement

The authors wish to thank Martin Berggren for valuable discussions about the VFA approach and for pointing out the reference [10].

Bibliography

[1] Burger, M., Pinnau, R.: Fast optimal design of semiconductor devices. SIAM Journal on Applied Mathematics **Vol. 64**(1) (2003) 108–126

[2] Hinze, M., Pinnau, R.: Optimal control of the drift-diffusion model for semiconductor devices. In K.H.Hoffmann, I.Lasiecka, G.Leugering, J.Sprekels, eds.: Optimal Control of Complex Structures. Volume 139 of ISNM. Birkhäuser (2001) 95–106

[3] Collis, S.S., Heinkenschloss, M.: Analysis of the Streamline Upwind/Petrov Galerkin method applied to the solution of optimal control problems. Technical Report CAAM TR02-01, CAAM, Rice University (March 2002)

[4] Abraham, F., Behr, M., Heinkenschloss, M.: The effect of stabilization in finite element methods for the optimal boundary control of the oseen equations. Finite Elements in Analysis and Design **41** (2004) 229–251

[5] Ridzal, D., Bochev, P.: A comparative study of Galerkin and mixed Galerkin discretizations of the state equation in optimal control problems with applications to semiconductor modeling. Technical Report SAND2007, Sandia National Laboratories (2007)

[6] Babuška, I., Miller, A.: The post processing approach in the finite element method, part 1: calculation of displacements, stresses and other higher derivatives of the displacements. Internat. J. Numer. Methods. Engrg **Vol. 34** (1984) 1085–1109

[7] Brezzi, F., Hughes, T.J.R., Süli, E.: Variational approximation of flux in conforming finite element methods for elliptic partial differential equations: a model problem. Rend. Mat. Acc. Lincei **ser. 9, vol. 12** (2001) 159–166

[8] Wheeler, J.A.: Simulation of heat transfer from a warm pipeline buried in permafrost. In: Proceedings of the 74th National Meeting of the American Institute of Chemical Engineers. (1973)

[9] Wheeler, M.F.: A Galerkin procedure for estimating the flux for two-point boundary value problems. SIAM Journal on Numerical Analysis **Vol. 11**(No. 4) (1974) 764–768

[10] Berggren, M., Glowinski, R., Lions, J.L.: A computational approach to controllability issues for flow- related models. (II): Control of two-dimensional, linear advection- diffusion and Stokes models. INJCF **Vol. 6**(4) (1996) 253–274

[11] Brezzi, F., Fortin, M.: Mixed and Hybrid Finite Element Methods. Springer, Berlin (1991)